

MODELO ESTOCASTICO PARA LA PREDICCIÓN DE LA VELOCIDAD DEL VIENTO

por E. Hernández y R. García
Meteorología Dinámica
Departamento de Física del Aire y Geofísica
Universidad Complutense, Madrid

Introducción.

Una de las variables meteorológicas cuya predicción resulta más dificultosa es el vector viento en superficie. Dada la gran influencia que tienen sobre su comportamiento los factores locales que se superponen a los patrones atmosféricos de comportamiento a meso y macroescala, la previsión ha sido tratada recientemente desde la óptica de los modelos estocásticos. En ellos se considera cada valor observado como la manifestación de un proceso estocástico generador de un conjunto de valores. De esta manera se puede determinar el comportamiento y evolución es futura de una serie temporal sobre las bases de su propio pasado y de un término aleatorio denominado ruido. El punto crucial de la predicción consiste en la identificación del filtro existente entre la variable y la serie de ruido, lo cual se lleva a cabo a través de las funciones de autocorrelación y autocorrelación parcial. En el campo meteorológico y ambiental cabe citar los trabajos de Tiao, Chock y Finzi.

Por lo que respecta a la previsión del viento, ésta ha sido abordada mediante pro-

cedimientos de Markov (Bacon et al.) y autorregresivos (Mage, Bonivento et al.) sobre las bases horarias con relativo éxito; en cambio, si se emplean valores medios diarios las dificultades son mucho mayores al presentarse menor persistencia y mayor variabilidad, por ende el acercamiento conseguido en la modelización es más deficiente (Luna et al., Bacci et al.).

La mayor parte de estos modelos han sido evaluados, como se mencionó anteriormente, sobre la base de técnicas de análisis de series temporales, sin considerar explícitamente la información que día a día proporcionan las variables meteorológicas. Ello hace que resulte de primordial interés la investigación de métodos que hagan factible la incorporación de tal aporte. De este modo surge la evaluación de filtros multivariantes, especialmente ARMAX (Young et al.) y funciones de transferencia (Box y Jenkins) que requieren registros simultáneos de un conjunto de variables meteorológicas susceptibles de ejercer una función rectora sobre el viento. Sin embargo, surge una alternativa a este tipo de modelos, considerando que el filtro interpuesto

entre el ruido y la velocidad es variable con la situación meteorológica reinante en el día considerado. De esta manera se puede aprovechar cualitativamente el caudal de conocimientos meteorológicos, haciendo los modelos más versátiles y adaptativos.

El objeto de este trabajo consiste en la modelización univariante de la velocidad media diaria del viento, tanto mediante técnicas de análisis de series temporales con filtros constantes como variables, prestando especial interés a la influencia que ejerce la introducción de la información meteorológica en los resultados.

Bases teóricas

Modelos ARMA con coeficientes constantes

De entre los distintos modelos estocásticos estacionarios se ha optado por recurrir a aquéllos cuya base está constituida por el tratamiento de series temporales, ya que han dado resultados satisfactorios en otras circunstancias (Tao et al., Chock et al.). En particular, se centró la atención sobre los modelos autorregresivos de media móvil (Box y Jenkins). En ellos interviene exclusivamente la información contenida en la propia serie temporal, es decir, son univariantes, lo que supone una gran cortapisa, pues, se desperdician las aportaciones que pueda proporcionar el conocimiento de variables relacionadas con la dependiente. Sin embargo, la determinación del orden autorregresivo de los procesos pondrá de manifiesto interesantes aspectos cualitativos y cuantitativos sobre la estructura de la persistencia.

Aunque su base es ciertamente empírica, poseen una estructura prefijada que se debe identificar y unos parámetros a cuya estimación hay que proceder. Consecuentemente, esto permite considerarlos intermedios entre

los determinísticos y los puramente empíricos, los cuales carecen de bases teóricas que justifiquen las relaciones entre las distintas variables que intervienen en los diversos fenómenos.

El fundamento en que se basan es el hecho, constatado por Youle, de que las series temporales con un cierto grado de persistencia se pueden generar a partir de la señal con que un filtro responde a una secuencia temporal aleatoria conocida como ruido. Puesto que las series a estudiar presentan generalmente una gran componente persistente, se procederá a una modelización univariante basándose en el hecho anteriormente reseñado. Dentro de los filtros que se pueden aplicar a este tipo de series, se centrará la atención en los lineales, cuya respuesta en un cierto instante, t , viene determinada por la suma ponderada de los valores que ha tomado previamente el ruido, es decir, que

$$Z_t = \mu + \Psi(B) a_t \quad (1)$$

donde:

Z_t es la serie temporal observada

a_t es el ruido

$\Psi(B)$ es un polinomio representativo del filtro que transforma a_t en Z_t , siendo la variable B el operador denominado de desplazamiento regresivo, es decir, que transforma cada valor de la serie en el inmediato anterior y, finalmente, μ es el parámetro que en ciertos casos coincide con la media y, en general, representa el nivel de la serie.

Una gran mayoría de las variables que aparecen en la realidad se presentan a través de series estacionarias, por lo que será interesante conocer las condiciones que debe cumplir $\Psi(B)$ para que la serie generada sea estacionaria. En efecto, si se tiene en cuenta que $\Psi(B)$ se puede considerar como función generatriz de los coeficientes Ψ_j tomando B

como una variable falsa cuya potencia j -enésima es el coeficiente Ψ_j , la condición de estacionariedad es que la serie $\Psi(B)$ converja cuando $|B| \leq 1$.

Asimismo, si en la expresión (1) el valor de la media es cero se tiene que, para determinar los valores de a_t en función de Z_t se deberá poder escribir

$$a_t = \Psi(B)^{-1} Z_t \quad (2)$$

o lo que es equivalente (Box y Jenkins)

$$a_t = -\Psi Z_{t-1} - \Psi^2 Z_{t-2} - \dots - \Psi^k Z_{t-k} + Z_t$$

Está claro que si Z_t toma un valor finito, la serie deberá converger, y será preciso que esté asegurada la invertibilidad para que se garantice la aceptabilidad de los valores de Z_t obtenidos.

A partir de la expresión (2) se obtiene, cuando existe $\Psi^{-1}(B)$,

$$\Pi(B) Z_t = a_t \quad (3)$$

es decir, que el polinomio $\Pi(B) = \Psi^{-1}(B)$ deberá converger para que (3) tenga sentido. La condición analítica de estacionariedad viene expresada por la convergencia de los coeficientes del polinomio $\Pi(B)$, tanto en el interior como en la periferia del círculo unidad. Es de destacar que en los modelos en los cuales el número de términos del polinomio $\Pi(B)$ es finito, el proceso de identificación y elaboración se ve muy facilitado. Por otra parte, las condiciones de invertibilidad se cumplen de manera inmediata, ya que al ser un número finito de términos, la serie converge sin imponer restricción alguna. La estacionariedad viene garantizada cuando las raíces

de la ecuación característica $\Pi(B) = 0$ se encuentran fuera del círculo unidad. Si el número de términos del polinomio es p , este será autorregresivo de orden p , con lo cual los valores de Z_t se pueden expresar como combinación lineal de los p valores previos de la propia serie temporal más un término de ruido, es decir, cada valor se autorregresiona sobre valores anteriores de sí mismo.

Cuando el polinomio $\Psi(B)$ posee un número finito de términos, q , el modelo es de media móvil de orden q , es decir, Z_t se puede expresar como una combinación lineal de un número finito de impulsos, a_t , del ruido. Por su propia naturaleza, estas series son siempre estacionarias, al tener $\Psi(B)$ un número finito de términos. De una manera análoga a las consideraciones anteriores, las series de este tipo son invertibles cuando las raíces de la ecuación característica $\Psi(B) = 0$ no están contenidas dentro del círculo unidad, ni en la circunferencia que lo limita.

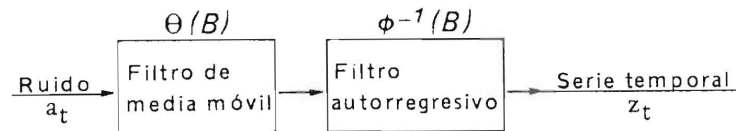
Combinando los modelos anteriores de manera que el valor de la serie se pueda expresar como función lineal de, p , valores previos de sí mismo y, q , del ruido se obtienen los modelos autorregresivos de media móvil de orden (p, q) . Las condiciones de invertibilidad y estacionariedad para estos modelos, vienen expresadas por la restricción de que las soluciones de las correspondientes ecuaciones características sean exteriores al círculo unidad.

Modelos con coeficientes variables

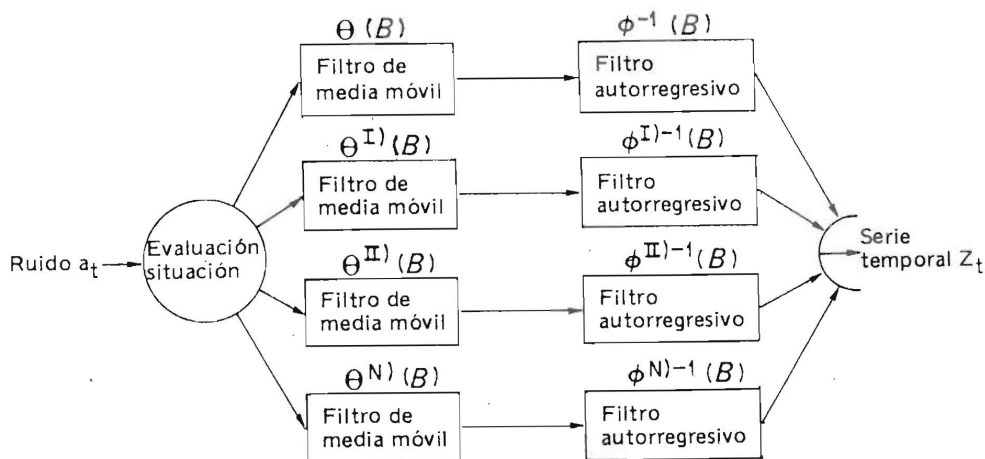
Como se puede deducir de todo lo anterior, las variables externas a la propia serie no influyen explícitamente sobre el comportamiento de la serie, sino que están englobadas en el término de ruido. Además, una de las restricciones más importantes de estos modelos es la constancia del filtro interpuesto en-

tre las variables consideradas. Para soslayarla se debe recurrir a técnicas en las cuales no se centra la homogeneidad de ciertos parámetros que hagan variar dicho filtro.

Así, mientras que el proceso de los modelos autorregresivos de media móvil se puede resumir como sigue:



el seguido en los modelos de filtro variable sería:



Así, la mayor versatilidad y adaptabilidad de estos modelos lleva aparejada un proceso de estimación más laborioso y complejo, tanto porque no se deben fijar primeramente las características definidoras de las situaciones, como por la diversidad de filtros que surgen.

La elección práctica de un modelo u otro dependerá en cada caso de los requerimientos que deba cumplir el modelo.

Modelos estocásticos

Para la elaboración de los modelos consecuentes con la filosofía estocástica expuesta en el apartado anterior, se ha tomado como

base de datos las observaciones de la velocidad media diaria del viento en el observatorio de Retiro (Madrid), durante el periodo comprendido entre el 1 de octubre de 1979 y el 31 de marzo de 1980.

Previamente a la identificación de cualquier modelo ARMA, se ha procedido a investigar las características que ha de satisfacer la serie para que éstos sean aplicables. En primer lugar, se ha verificado el carácter homoscedástico de la serie, para ello se construyó el correspondiente diagrama rango-media de la serie original representado en la figura 1. Puede observarse claramente que la serie no cumple tal condición por lo que se hace imprescindible una transformación. Después de la aplicación de varias de ellas, se encontró que la más

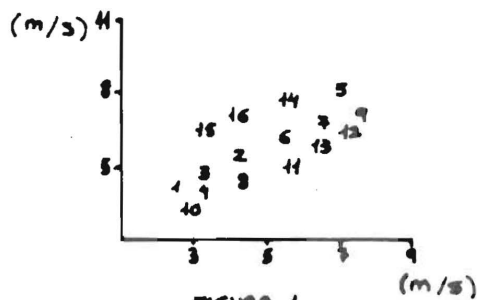


FIGURA 1
Diagrama rango-media de la serie V_t .

adecuada para inducir esta propiedad era la logarítmica, tal y como puede comprobarse en la figura 2. En consecuencia, la serie a modelizar está constituida por los valores del logaritmo de la velocidad media diaria. Dadas las características de la transformación, las calmas se han asimilado a los valores correspondientes al límite de detección del instrumento de medida.

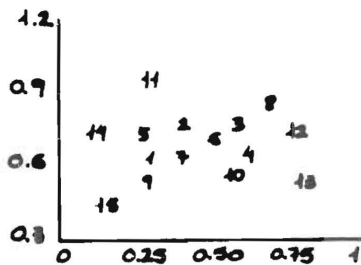


FIGURA 2
Diagrama rango-media de la serie $\lg V_t$.

A partir de estos valores se ha calculado la función de autocorrelación que se halla representada en la figura 3. Puede observarse que la serie es estacionaria y, por ello, se puede proceder sin ningún otro tipo de transformación a la identificación del filtro. Esta se ha realizado a partir, tanto de la propia función de autocorrelación como de la autocorrelación parcial (ver figura 4). Analizando ambas se ha encontrado que el modelo apropiado para la serie resulta ser autorregresivo de primer orden, con un único parámetro a estimar, cuyo valor es 0.55, después de llevar a cabo

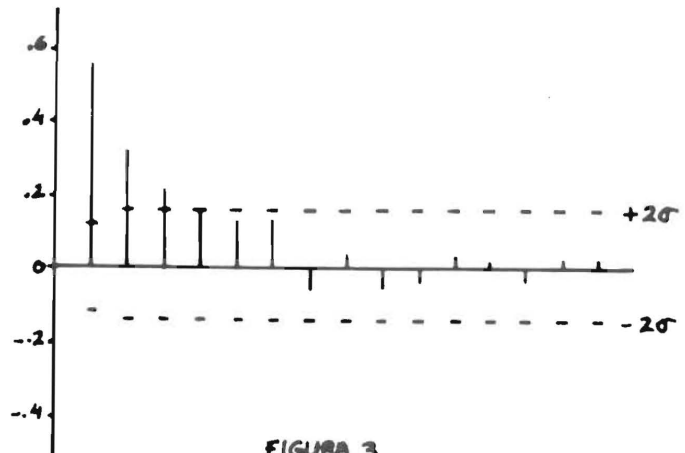


FIGURA 3
Función de autocorrelación de la serie $\lg V_t$.

un proceso estimativo con retrovisión. Por tanto, la expresión final del modelo es

$$\lg V_t = 0.55 \lg V_{t-1} + a_t \quad (4)$$

donde a_t representa el término de ruido.

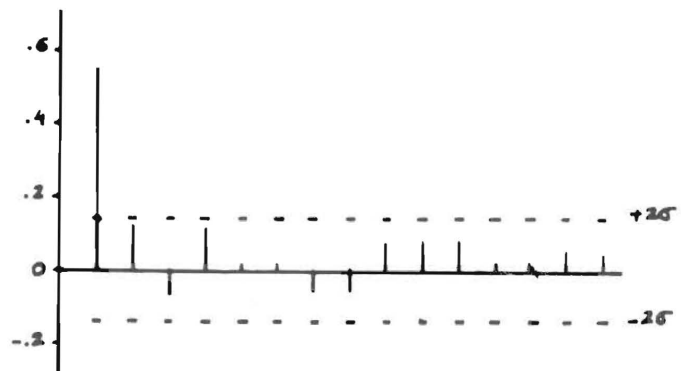


FIGURA 4
Función de autocorrelación parcial de la serie $\lg V_t$.

De (4) se deduce que la persistencia que presenta el viento es simple y que las influencias de los agentes externos únicamente son significativos durante el día para el cual se realiza la predicción.

La calidad del modelo se evaluó sobre la serie sin transformar, ya que ésta es la que in-

teresa predecir. La evaluación de estos resultados se realizó sobre el periodo comprendido entre el 1 de octubre de 1980 y el 31 de marzo de 1981, obteniéndose una correlación de 0.53 entre los valores predichos y observados, que resulta significativa al nivel del 95 por ciento, siendo además, apreciablemente superior a la obtenida con persistencia markoviana, cuyo valor es de 0.4.

Determinadas así las características del modelo ARMA y su capacidad predictora, se ha procedido a la elaboración de un modelo de filtro variable. El primer problema que se presenta es la elección del conjunto de variables y parámetros capaz de discriminar en cada momento el filtro "ad hoc". Como quiera que el viento viene muy influenciado por el patrón sinóptico actuante sobre el punto de muestreo, parece claro que ésta será una de las influencias determinantes sobre cualquier modelo que intente explicar su comportamiento. Por ello, se procedió a discriminar los días según los criterios determinantes de las situaciones sinópticas típicas de la Península (Hernández et al.). De acuerdo con esta clasificación se ha encontrado que el filtro múltiple interpuesto entre la serie temporal de viento y la de ruido posee seis componenets que corresponden a las situaciones 3, 4, 5, 7, 8 y R, respectivamente. (R engloba aquellos días que no corresponden a las situaciones mencionadas explícitamente). De acuerdo con esto se ha evaluado el conjunto de parámetros asociados a los diferentes filtros. Se debe señalar que la estimación se ha particularizado para cada una de las situaciones sin mantener la estructura de serie temporal. En la tabla I se presenta un resumen de los resultados obtenidos.

En la primera columna se indica el tipo de modelo para cada situación, en la siguiente se da el valor del parámetro asociado. La otras dos corresponden al coeficiente de correlación entre los valores predichos y obser-

Situación	Modelo	Parámetro asociado	r	σ_r/σ
3	AR(1)	0.53	0.61	0.62
4	"	0.59	0.54	0.71
5	"	0.23	0.56	0.63
7	"	0.49	0.58	0.66
8	"	0.52	0.60	0.64
R	"	0.70	0.60	0.64
Global	"		0.66	0.71

TABLA 1.— Resultados del modelo con filtro variable

vados y al cociente entre la desviación típica de los residuos y la serie original respectivamente.

En primer lugar cabe destacar la notable mejora que supone la introducción de las categorías sinópticas en el modelo, aumentando la varianza explicada en un 15 por ciento con respecto al modelo con un único filtro y en 27 por ciento con respecto a la persistencia. Todos los coeficientes de correlación obtenidos son significativos al nivel del 95 por ciento. Es de resaltar que, de acuerdo con las dos últimas columnas, el nivel de predicción para cada una de las situaciones es muy similar y no se presentan diferencias significativas. Por otra parte, el modelo identificado coincide con el de filtro único y no varía de una situación a otra, aunque los valores de los parámetros reflejan el distinto grado de persistencia simple de cada una de ellas. Los resultados que se han valorado se han obtenido en tiempo real, obedeciendo la clasificación de las categorías a los mapas sinópticos predichos por el I.N.M.

Conclusiones

A la vista de los anteriores resultados cabe destacar que el empleo de los modelos estocásticos univariantes, los más sencillos de entre ellos, permite alcanzar un nivel de previsibilidad significativo únicamente con los valores de la propia serie y de la información

sinóptica diaria habitual.

Sin tener en cuenta explícitamente ninguna información meteorológica, el modelo resultante es AR(1) y da lugar a un coeficiente de correlación entre los valores observados y los del modelo de 0.53, significativo al nivel del 95 por ciento. Hay que señalar que, a pesar de que se cuenta exclusivamente con una variable, se mejora apreciablemente la predicción que proporciona la persistencia markoviana.

Se ha puesto de manifiesto el papel crucial que juegan los patrones sinópticos en el modelo de predicción, mejorando la calidad de la misma significativamente. Los modelos correspondientes a las diferentes categorías son todos de carácter autorregresivo de primer orden. Los parámetros de dichos modelos presentan una gran variabilidad, pero no así los resultados que producen, los cuales se pueden resumir en un coeficiente de correlación de 0.66 y un cociente σ_r/σ de 0.71.

ABSTRACT

On the basis of daily average wind speed data, a forecasting stochastic model has been built. Since one of the main influences in wind speed is the synoptic pattern over the measurement point, a new model has been elaborated taking into account the meteorological information. The model obtained with the meteorological input gives a better performance than the univariate one.

BIBLIOGRAFIA

- Bacci, G. and Finzi, G. *A statistical predictor of surface wind speed in Milan city*. WMO-RA-VI Meeting on forecasting of conditions leading to high levels of air pollution occurrence. Leningrad, USSR, 17-20 nov. (1980).
- Bacon, J.W. and Herdenson, B. *The generation of hourly average wind vectors using a Markov process*. *Atm. Pollut.* Elsevier, Amsterdam (1980).
- Bonivento, C. et al. *Real time prediction of local wind by means of stochastic models*. *Proceeding of the 14th Int. Coll. París*, 105-108 (1980).
- Box, J.E.P. and Jenkins, J.W. *Time series analysis: modeling and forecasting*. Holden Day, San Francisco (1976).
- Chock, D.P. et al. *Time series analysis of Riverside, California, air quality data*. *Atm. Env.* 9. 978-989 (1975).
- Finzi, G. et al. *Modelling and forecast of the Dosage population product in Venice*. IFAC Symposium on Environmental Systems, Planning, Design and Control, Kyoto, Japan (1977).
- Hernández, E. y Soler, A. *Situaciones locales típicas: parámetros o variables más características*. *Revista de Geofísica*, XXXVI, 1 (1978).
- Luna, R.E. and Church, H.W. *Estimation of long term concentration using an universal wind speed distribution*. *J. App. Meteor.* 13. 910-916 (1974).
- Mage, D.T. *Frequency distribution of hourly wind speed measurements*. *Atm. Env.* 14. 367-374 (1980).
- Tiao, G.C. et al. *A statistical analysis of the Los Angeles ambient carbon monoxide data 1955-1972*. *J. Air. Pollut. Control. Ass.* 25. 1.129-1.136 (1975).
- Young, P. and Whitehead *A recursive approach to time series analysis for multivariate systems*. *Int. J. Control.* 25. 457-482 (1977).
- Youle, G.U. *On a method of investigating periodicities in disturbed series, with special reference to Wolfers sunspot numbers*. *Phil. Trans. A* 226. 267-273 (1927).